Homework 7 Solutions - 165 Total Points

STAT 201-502

Lecture 11, 12, & 13 Material

1. Studies that compare treatments for chronic medical conditions such as headaches can use the same subjects for each treatment. With a crossover design, each person crosses over from using one treatment to another during the study. One such study considered a drug (a pill called Sumatriptan) for treating migraine headaches in children (ML Hamalainen et al, *Neurology*, 48, 1997). The study observed each of 30 children at tow times when he/she had a migraine headache. The child received the drug at one time and a placebo at the other time. The order of treatment was randomized and the study was double-blind. For each child, the response was whether the drug or the placebo provided better pain relief. Let p denote the proportion of children having better pain relief with the drug, in the population of children who suffer periodically from migraine headaches. Can you conclude that $p > 0.50$, with more than half of the population getting better with pain relief with the drug, or that $p < .50$, with less tha half getting better pain relief with the drug (ie, the placebo being better)? Of the 30 children, 22 had more pain relief with the drug and 8 had more pain relief with the placebo.

   (a) For testing $H_o : p = 0.50 \; VS \; H_a : p \neq 0.50$, show that the test statistic $Z = 2.56$.
   $Z = \frac{\hat{p}-p_o}{\sqrt{\frac{p_o(1-p_o)}{n}}} = Z = \frac{0.733-0.50}{\sqrt{\frac{0.5(0.5)}{30}}} = \frac{.233}{.091} = 2.56$

   (b) Show that the P-value is 0.0104. Interpret.
   p-value=$P(Z > 2.56) + P(Z < -2.56) = 2P(Z < -2.56) = 2(0.0052) = 0.0104$. This p-value is small so you reject the null hypothesis. The study shows that the proportion of children that get pain relief from the drug is greater than one half.

   (c) Check the assumptions needed for this test and discuss the limitations due to using a convenience sample.
   It's a convenience sample so it's not random.
   $n(p_o) = 30(.5) = 15$ and $n(1 - p_o) = 30(.5) = 15$ so the expected number of successes and failures are large enough to assume normality.

2. According to an exit poll in the 2000 New York senatorial election, 55.7% of the sample of size 2232 reported voting for Hillary Clinton. Is this enough evidence to predict who own? Test that the population proportion who voted for Clinton was 0.50 against the alternative that it differed from 0.50. Answer by

   (a) Identifying the variable and parameter, and defining the notation.
   The variable is whether a person voted for Hillary Clinton.
   The parameter we are interested in is p, the proportion of people in the population that voted for Hillary Clinton.

   (b) Stating the hypotheses and checking assumptions for a large-sample test. $H_o : p = p_o\ VS.\ H_a : p \neq p_o$
   $n(p_o) = 2232(.5) = 1116$ and $n(1 - p_o) = 2232(.5) = 1116$

   (c) Finding the test statistic for the test.
   $Z = \frac{\hat{p}-p_o}{\sqrt{\frac{p_o(1-p_o)}{n}}} = Z = \frac{0.557-0.5}{\sqrt{\frac{.5(.5)}{2232}}} = \frac{.057}{.01} = 5.7$

   (d) Reporting the P-value and interpreting it. p-value=$P(Z > 5.7) + P(Z < -5.7) = 2P(Z < -5.7) = 2(0) = 0$. This p-value is small so you reject the null hypothesis. The true number of people who voted from Hillary Clinton is actually greater than one half.

   (e) Explaining how to make a decision for the significance level of 0.05. Since the p-value is less than 0.05 we reject the null hypothesis as stated in the above section.

3. A study has a random sample of 20 subjects. The test statistic for testing $H_o : \mu = 100$ is $Z = 2.41$. Find the approximate P-value for the alternative

   (a) $H_a : \mu \neq 100$
   $P(Z > 2.41) + P(Z < -2.41) = 2P(Z < -2.41) = 2(0.008) = 0.016$

   (b) $H_a : \mu > 100$
   $P(Z > 2.41) = 1 - P(Z < 2.41) = 1 - 0.992 = 0.008$

   (c) $H_a : \mu < 100$
   $P(Z < 2.41) = 0.008$

4. Do women tend to spend more time6-38 on housework than men? If so, how much more? Based on data from the National Survey of Families and Households, one study (Souch and Spitze, *American Sociological Review*, 59, 1994) reported the results in the table for the number of hours spend in housework perweek.

$$P(Z > -2.41) + P(Z < -2.41) = 2P(Z < -2.41) = 2(0.008) = 0.016$$

Gender	**Sample Size**	**Mean**	**Standard Deviation**
Women	6764	32.6	18.2
Men	4252	18.1	12.9

(a) Show that this study estimated that, on the average, women spend 14.5 more hours a week on housework than men.
$32.6 - 18.1 = 14.5$

(b) Find the standard deviation for comparing the means. What factor causes the standard deviation to be small compared to the sample standard deviations from the two groups?
$\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{18.2^2}{6764} + \frac{12.9^2}{4552}} = \sqrt{0.049 + 0.037} = \sqrt{0.086} = 0.293.$
The standard deviation of the difference is smaller because (1) the values of differences are smaller and (2) we are looking at the average of the two standard deviations of the two groups and standard deviations of averages are always smaller.

5. A crossover study of 13 children suffering from asthma (*Clinical and Experimental Allergy*, 20, 1990) compared single inhaled doses of formoterol(F) and salbutamol(S). The outcome measured was the child's peak expiratory flow(PEF) eight hours following treatment. The data on PEF follow:

Child	**F**	**S**	**Child**	**F**	**S**	**Child**	**F**	**S**
1	310	270	6	370	300	10	250	210
2	385	370	7	410	390	11	380	350
3	400	310	8	320	290	12	340	260
4	310	260	9	330	365	13	220	90
5	410	380						

Let μ denote the population mean of the difference between the PEF values for the F and S treatments.The variance of the difference is 44.65.

(a) Calculate the 13 differences between the two scores.
40 15 90 50 30 70 20 30 -35 40 30 80 130

(b) Carry out the five steps of the significance test for a mean of the difference scores using $H_o : \mu = 0 \; VS \; H_a : \mu \neq 0$
(1)Assumptions: we don't know if it's random but we hope so and $n < 30$ so we can't assume normality. The hypothesis test may not be valid.
(2)Hypotheses: $H_o : \mu = 0 \; VS \; H_a : \mu \neq 0$ where μ is the mean of the difference scores.
(3)Test statistic: First we need to find the average of the sample differences $\bar{X} = 45.38$.
$Z = \frac{\bar{X}-\mu_o}{\frac{\sigma}{\sqrt{n}}} = \frac{45.38-0}{\frac{\sqrt{44.65}}{\sqrt{13}}} = \frac{45.38}{1.85} = 24.53$
(4)P-value $= P(Z > 24.53) + P(Z < -24.53) = 2P(Z < -24.53) = 2(0) = 0$
(5)P-value is so small that we reject the null hypothesis. We should not completely trust this conclusion since our assumptions for the test were not met.

6. Chelation is an alternative therapy for heart disease that uses repeated intravenous administration of human-made amino acid in combination with oral vitamins and minerals. Practitioners believe it removes calcium deposits from buildup in arteries. However, the evidence for a positive effect is anecdotal or comes from nonrandomized, uncontrolled studies. A double blind random clinical trial comparing chelation to placebo used a treadmil test in which the repose was the length of time until a subject experienced ischaemia (lack of blood flow and oxygen to the heat muscles).

   (a) After 27 weeks of treatment, the sample mean time for the chelation group was 652 seconds. A 95% confidence interval for the population mean for chelation minus the population mean for placebo was -53 to 36 seconds. Explain how to interpret the confidence interval.
   In 95% of the confidence intervals from all possible samples, the true mean difference would fall in this interval.

   (b) A significance test comparing the means had P-value = 0.69. Specify the hypotheses for this test, which is two-sided.
   $H_o : \mu = 0\ VS.\ H_a : \mu \neq 0$
   The p-value is so high that we can not reject the null hypothesis. There is no difference between the chelation therapy and the placebo effect.

   (c) The authors concluded from the test that, "There is no evidence to support a beneficial effect of chelation therapy." Explain how this conclusion agrees with inference based on the values in the confidence interval.
   We can also test that there is no true difference between the two therapies seeing if 0 is a possible value for μ in our confidence interval in part (a). Since the confidence interval contains 0 then we can say that there is no evidence to support a difference in time between chelation therapy and the placebo.

7. Do textbook problems 6.40(p418), 6.42(p419), 6.55(p421), 6.56(p421) and 8.22(p553)

   6.40
   (a) do not reject
   (b) do not reject
   (c) do not reject both because the p-value is greater than the significance level (α) and we only reject when it is less than the significance level.

   6.42
   (a)p-value for $H_a : \mu > \mu_o$ is $P(Z > z) = P(Z > 1.6) = 1 - P(Z < z) = 1 - .9452 = .0548$
   (b) p-value for $H_a : \mu < \mu_o$ is $P(Z < z) = P(Z < 1.6) = .9452$
   (c) p-value for $H_a : \mu \neq \mu_o$ is $2 \times P(Z > |z|) = 2 \times P(Z > 1.6) = 2 \times .0548 = .1096$

6.55

(a)$\bar{X} = 132.2$, $n = 25$ and $\sigma = 30$ let C=.95 therefor .07(.031)=2.26

P-value $= P(Z > 2.26) + P(Z < -2.26) = 2P(Z < -2.26) = 2(0.0119) = 0.0238$. P-value is less than 0.05 so reject H_o. There is evidence showing that Guatemalan women have higher measurement that the healthy level.

(c) $\bar{X} \pm Z^*(\frac{\sigma}{\sqrt{n}} = 9.57 \pm 1.96(0.031)$.

$Z = \frac{\bar{X}-\mu_o}{\frac{\sigma}{\sqrt{n}}} = \frac{132.2-115}{\frac{30}{25}} = 2.87$

$H_a : \mu > 115$ so the p-value $= P(Z > 2.87) = 1 - P(Z < 2.87) = 1 - .9979) = .0021$

For a α=.05 significance level reject H_o because $Z = 2.87 > 1.96 = Z^*$ or because p-value=.0021 ¡ .05=α. We reject H_o means that there is enough evidence to conclude these older students have a better attitude toward school in terms of their score on the psychological test.

(b) It is important that the 25 students were a random sample from all students aged 30 or older which is a crucial assumption. The less crutial assumption is a normally distribution but this is not as crutial as long as there are no outliers and very little skewness.

6.56

(a) $H_o : \mu = 9.5\ VS.\ H_a : \mu \neq 9.5$

(b) $Z = \frac{\bar{X}-\mu_o}{\frac{\sigma}{\sqrt{n}}} = \frac{9.57-9.5}{\frac{0.4}{\sqrt{160}}} = \frac{0.07}{0.031} = 2.26$

P-value $= P(Z > 2.26) + P(Z < -2.26) = 2P(Z < -2.26) = 2(0.0119) = 0.0238$. P-value is less than 0.05 so reject H_o. There is evidence showing that Guatemalan women have higher measurement that the healthy level.

(c) $\bar{X} \pm Z^*(\frac{\sigma}{\sqrt{n}} = 9.57 \pm 7.96(0.031) = (9.51, 9.63)$. The confidence interval for the true level for Guatemalan women is extremely close to 9.5 so it's not necessarily practical to reject the null hypothesis under these circumstances.

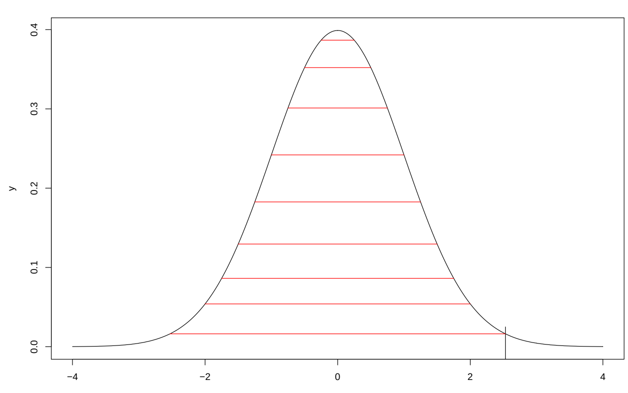
8.22

(a) $H_o : p \geq 0.5\ VS.\ H_a : p < 0.5$

$Z = \frac{\hat{p}-p_o}{\sqrt{\frac{p_o(1-p_o)}{n}}} = Z = \frac{0.7-0.50}{\sqrt{\frac{0.5(0.5)}{40}}} = \frac{0.2}{0.29} = 2.53$

p-value=$P(Z < 2.53) = 0.9943$

(b)



(c)Do not reject H_o. The majority of people prefer fresh-brewed coffee.